Investigaciones Geográficas, Boletín del Instituto de Geografía, UNAM
ISSN 0188-4611, Núm. 82, 2013, pp. 118-130

Evaluación y comparación de métodos de interpolación determinísticos y probabilísticos para la generación de modelos digitales de elevación

Recibido: 19 de junio de 2012. Aceptado en versión final: 13 de noviembre de 2012.

Cutberto Uriel Paredes Hernández*
Wilver Enrique Salinas Castillo*
Xicoténcatl Martínez Becerra**
Sergio Bernardo Jiménez Hernández***

Resumen. Los modelos digitales de elevación (MDE) son la forma más común y eficaz de representar la superficie terrestre. Desafortunadamente, al igual que la mayoría de la información geográfica en formato digital, los MDE contienen errores de forma inherente, por lo que es necesario evaluar la magnitud de dichos errores e informar de su existencia a los usuarios de MDE. Uno de los factores que introduce errores en los MDE es el método de interpolación y diferentes estudios previos han presentado conclusiones que discrepan entre sí en cuanto a la idoneidad de diferentes métodos de interpolación para la generación de MDE. Con el objetivo de contribuir a dicha discusión y de mejorar el entendimiento del error introducido por métodos de interpolación en los MDE, en este trabajo se evalúa y compara la precisión de MDE generados utilizando métodos de interpolación determinísticos (IDW, ANUDEM) y probabilísticos (OK, KT). Los resultados sugieren que KT genera MDE más precisos que el resto de los métodos de interpolación aquí evaluados y sin la presencia de errores sistemáticos, como ocurre en los MDE generados usando IDW, ANUDEM y OK. Dichos resultados concuerdan con algunos estudios previos, pero discrepan al mismo tiempo con otros estudios similares.

Palabras clave. Modelos Digitales de Elevación (MDE), interpolación, IDW, ANUDEM, kriging.

Evaluation and comparison of deterministic and probabilistic interpolation methods for digital elevation model generation

Abstract. Digital elevation models (DEM) are the most common and effective representation of the land surface. Unfortunately, similarly to most digital geographic information, DEM inherently contain errors, which magnitude should be assessed and informed to DEM users. One factor that introduces errors in DEM is the interpolation method and several previous studies have presented findings that disagree among them regarding the suitability of several interpolation methods for DEM generation. With the aim of contributing to this discussion and of improving the understanding of

* Unidad de Geomática, Facultad de Ingeniería y Ciencias, Universidad Autónoma de Tamaulipas, Centro Universitario Victoria, 87149, Ciudad Victoria, Tamaulipas, México. E-mail: cparedes@uat.edu.mx; wsalinas@uat.edu.mx
** GeoExpert S.C., Normal Superior 182, Col. Adolfo López Mateos 2A Etapa, 87025, Ciudad Victoria, Tamaulipas, México. E-mail: xicotencatl_mtz_b@hotmail.com
*** Centro de Investigación y Desarrollo en Ingeniería Portuaria, Marítima y Costera, Facultad de Ingeniería "Arturo Narro Siller", Universidad Autónoma de Tamaulipas, CIDIPORT-UAT Tampico, Tamaulipas, México. E-mail: sjimenez@uat.edu.mx

errors introduced in DEM by interpolation methods, in this paper we evaluate and compare the accuracy of DEM generated using both deterministic (IDW, ANUDEM) and probabilistic (OK, KT) interpolation methods. Our results suggest that KT produces more accurate DEM than the other interpolation methods evaluated in this paper and that it produces DEM without systematic errors, which occur in DEM generated using IDW, ANUDEM and OK. These results are consistent with some previous studies, but at the same time disagree with other similar studies.

Key words: Digital Elevation Models (DEM), interpolation, IDW, ANUDEM, kriging.

INTRODUCCIÓN

La superficie terrestre

La superficie terrestre es un fenómeno continuo que puede ser medido y descrito en cualquier punto sobre la misma con base en un conjunto de variables (Longley *et al*., 2005). Entre dichas variables, la elevación es una de las variables más estudiadas, ya que es utilizada en un amplio rango de aplicaciones científicas y civiles (INEGI, 2002; Fisher y Tate, 2006), y se define como la distancia vertical desde un punto sobre la superficie terrestre hasta un nivel de referencia arbitrario, como puede ser el nivel medio del mar (definido como datum vertical), (Anderson y Mikhail, 1998).

Debido a la naturaleza de la superficie terrestre, la elevación es continua y debe ser representada digitalmente como tal. Desafortunadamente, debido a limitantes técnicas y económicas, no es factible medir y almacenar la elevación de cada punto sobre la superficie terrestre (Kumler, 1994). Aun considerando nuevas tecnologías como Lidar (*Li*ght *D*etection and *R*anging), que brinda la posibilidad de medir la elevación de un considerable número de puntos sobre el terreno en un corto periodo de tiempo, no es viable representar la superficie terrestre utilizando todas las muestras de elevación (puntos sobre el terreno) disponibles y es necesario reducir el volumen de datos (Anderson *et al*., 2006). Por tanto, la elevación del terreno se representa digitalmente como una superficie continua mediante Modelos Digitales de Elevación (MDE) generados con base en un número finito de muestras de elevación (Li *et al*., 2005).

Modelos digitales de elevación (MDE)

Las muestras de elevación o puntos sobre el terreno no representan la superficie terrestre de manera continua por sí mismas, sin importar si su distribución espacial es regular o irregular (Kumler, 1994), y, en consecuencia, es necesario utilizar un formato que permita modelar la elevación como una superficie continua (Burrough y McDonell, 1998). De los diferentes formatos disponibles (curvas de nivel, redes irregulares de triángulos (TIN), celdas (Li *et al*., 2005)), el formato de celdas, o *raster*, es el más común entre los usuarios y es normalmente conocido como Modelo Digital de Elevación (MDE), (Fisher y Tate, 2006). Sin embargo, las curvas de nivel y las redes irregulares de triángulos pueden ser también consideradas como MDE si se utiliza la definición tradicional de MDE de Miller y Laflamme (1958): "representación numérica digital de la superficie terrestre". En este artículo se utiliza el término MDE para referirse únicamente al formato de celdas, como es común en la literatura científica.

Los MDE en formato de celdas representan el terreno como una superficie continua mediante la definición de una cuadrícula o malla regular sobre el área en estudio y el almacenamiento de un valor de elevación para cada celda de la cuadrícula (Burrough y McDonell, 1998). La definición de la cuadrícula depende directamente de la distribución espacial de las muestras de elevación. Si los puntos sobre el terreno disponibles se distribuyen espacialmente en forma de matriz, entonces es recomendable definir la malla con base en las muestras de elevación disponibles y almacenar la elevación directamente en cada celda. En este caso el tamaño de celda o resolución espacial está determinado por la distancia entre cada muestra de elevación (Li *et al*., 2005).

En caso contrario, si las muestras de elevación se distribuyen irregularmente sobre el área de estudio, es necesario definir el tamaño de celda de manera arbitraria, empírica (Hutchinson, 1996) o estadística (Hengl, 2006), y seleccionar un método de interpolación para generar un valor de elevación para cada celda de la malla tomando como entrada los puntos sobre el terreno disponibles (Li *et al*.,

2005). Desafortunadamente, diferentes métodos de interpolación calculan diferentes valores para una misma celda aun utilizando las mismas muestras de elevación como entrada (Lloyd y Atkinson, 2002), por lo que la selección del algoritmo de interpolación puede ser considerada como un factor que introduce errores en los MDE.

Errores por interpolación en los MDE

Los errores introducidos por el método de interpolación seleccionado pueden ser aleatorios, debido por ejemplo al redondeo numérico (USGS, 1997); sistemáticos, manifestados como patrones en los MDE (Wise, 2000); o inclusive garrafales (*blunders*), causados por alta variabilidad del terreno y la aplicación de métodos de interpolación sin restricciones (Smith *et al*., 2005). Los errores sistemáticos y garrafales pueden ser eliminados o reducidos mediante la aplicación de filtros estadísticos apropiados (López, 2000), por lo cual no es común observarlos en MDE comerciales u oficiales (Fisher y Tate, 2006). Sin embargo, los errores aleatorios no pueden ser filtrados, debido al poco entendimiento de su naturaleza y localización (Li *et al.*, 2005; Pérez y Mas, 2009), y en algunos casos la aplicación de filtros para la reducción/eliminación de errores sistemáticos y garrafales resulta en la adición de nuevos errores aleatorios en los MDE (Wise, 2000).

Estudios previos han sugerido que los principales factores que influyen en la naturaleza y magnitud de los errores introducidos en los MDE por el método de interpolación son: *a)* la cantidad de información (ej. número de muestras, correlación espacial) empleada por el método de interpolación para calcular la elevación en cada celda de la malla (Lloyd y Atkinson, 2006) y *b)* los rasgos topográficos del área geográfica representada por el MDE (Fisher y Tate, 2006). No obstante, no existe un consenso en cuanto a ninguno de estos dos factores. Por ejemplo, para la generación de MDE en áreas con relieve moderado, Lloyd y Atkinson (2002), Reuter *et al*. (2007) y Erdogan (2009) concluyeron que métodos probabilísticos que emplean gran cantidad de información para calcular la elevación de cada celda son más convenientes que métodos determinísticos. Sin embargo, Bishop y McBratney (2002), Su y Bork (2006) y Bater y Coops (2009) sugieren que métodos determinísticos producen MDE más precisos en áreas de relieve similar cuando una gran cantidad de muestras de elevación es utilizada.

En áreas con relieve predominantemente montañoso, los resultados de Reuter *et al.* (2007) sugieren un mejor desempeño de métodos determinísticos en este tipo de terreno. En contraste, Morillo *et al*. (2002) no encontraron diferencias significativas en la aplicación de métodos determinísticos y probabilísticos para la generación de MDE y lo atribuyeron a la alta densidad de datos disponibles para su área en estudio. De forma similar, Guo *et al*. (2010) sugieren que diferentes métodos de interpolación generan resultados similares, aunque establecen que métodos probabilísticos son más confiables en la generación de MDE en áreas con relieve montañoso. Por lo tanto, es difícil establecer la superioridad de ciertos métodos de interpolación sobre otros y es necesario ampliar el nivel de entendimiento de los mismos con base en nuevos estudios (Fisher y Tate, 2006; Wise, 2011).

Objetivo

Con el objetivo de contribuir a un mejor entendimiento de los errores introducidos por diferentes métodos de interpolación en la generación de MDE, en este estudio se presenta la evaluación y comparación de la precisión de MDE generados utilizando métodos de interpolación determinísticos y probabilísticos. La descripción de los métodos utilizados en este estudio se presenta a continuación; seguida de un caso de estudio que permite comparar la idoneidad de cada método para la generación de MDE y alcanzar una serie de conclusiones, presentadas al final de este estudio.

MÉTODOS DE INTERPOLACIÓN

Métodos determinísticos

La interpolación espacial es "un procedimiento que permite calcular el valor de una variable en una posición del espacio, conociendo los valores de esa variable en otras posiciones del espacio" (Bosque, 2000:375). Los diferentes métodos de interpo-

lación espacial se clasifican en dos grandes categorías: determinísticos y probabilísticos (Maune *et al.*, 2001). Los métodos de interpolación determinísticos calculan un valor para celda de la cuadrícula usando únicamente las propiedades físicas de las muestras de elevación (Goovaerts, 1997). El método de interpolación determinístico más comúnmente empleado es la ponderación en función inversa de la distancia o IDW (*Inverse Distance Weighting*; Wise, 2000). IDW se encuentra disponible en la mayoría de los programas de sistemas de información geográfica (SIG; Pérez y Mas, 2009) y su definición formal es (O'Sullivan y Unwin, 2003:228):

$$\hat{e}_c = \frac{1}{n}\sum_{m=1}^{n} p_{mc} e_m \qquad (1)$$

donde $\hat{e}_c$ es la elevación calculada para la celda c utilizando n muestras de elevación; e_m es el valor de las muestras de elevación; y p_{mc} es el peso (o influencia) entre 0 y 1 asignado a cada muestra de elevación, calculado como:

$$p_{mc} = \frac{1}{d_{mc}} \qquad (2)$$

donde d_{mc} es la distancia euclidiana entre la muestra de elevación m y el centro de la celda c de la malla. Opcionalmente, con el fin de mejorar resultados, es común utilizar un exponente a para disminuir la influencia de las muestras conforme se incrementa la distancia (Bosque, 2000:378):

$$p_{mc} = \frac{1}{d_{mc}^{a}} \qquad (3)$$

Otro método de interpolación determinístico ampliamente utilizado es el método ANUDEM, desarrollado por Hutchinson (1988) específicamente para la generación de MDE usando curvas adaptivas (*splines*) locales. ANUDEM define un modelo estadístico en el que las muestras de elevación z_i en las coordenadas x_i, y_i son el resultado de (Hutchinson, 2008:151):

$$z_i = f(x_i, y_i) + \varepsilon_i \quad (i = 1,\ldots,n) \qquad (4)$$

donde f es una función bivariada desconocida potencialmente suave de ubicación horizontal representada como una malla de diferencia finita; n el número de muestras de elevación; y ε_i es el error aleatorio con media 0 y desviación estándar dada por:

$$w_i = hs_i\sqrt{12} \qquad (5)$$

donde h es el tamaño de celda de la malla y s_i es la pendiente de la celda asociada con la muestra. La función f es estimada resolviendo la aproximación de la malla de diferencia finita a la función bivariada que minimiza:

$$\sum_{i=1}^{n}\left[\frac{z_i - f(x_i, y_i)}{w_i}\right]^2 + \Lambda J(f) \qquad (6)$$

donde $J(f)$ es la rugosidad del terreno en función de la primera y segunda derivada de f; y Λ es el parámetro positivo que suaviza dicha rugosidad. El parámetro Λ es elegido de forma que la suma de cuadrados de los residuales en la ecuación 6 sea igual a n, lo cual solo puede ser logrado mediante un método de interpolación iterativa para el cual la pendiente de cada celda se encuentre disponible. Es decir, el método ANUDEM genera un MDE de baja resolución e interpola de forma iterativa incrementando la resolución del mismo hasta alcanzar la solución de la ecuación 6 y el tamaño de celda especificado por el usuario (*Ibid.*).

Métodos probabilísticos

Los métodos de interpolación probabilísticos, en contraste con los métodos determinísticos, infieren una gran cantidad de información utilizando las muestras disponibles (Maune *et al.*, 2001). Los métodos de interpolación probabilísticos más comunes son los métodos geoestadísticos derivados de la forma genérica de Kriging (Goovaerts, 1997:126):

$$Z^*(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha(\mathbf{u})\,[Z(\mathbf{u}_\alpha) - m(\mathbf{u}_\alpha)] + m(\mathbf{u}) \qquad (7)$$

donde $Z^*(\mathbf{u})$ es el valor estimado para cada ubicación $\mathbf{u}$ utilizando $n(\mathbf{u})$ muestras; λ_α es el peso,

o ponderación, que cada muestra $Z(\mathbf{u}_\alpha)$ toma; y $m(\mathbf{u}_\alpha)$ y $m(\mathbf{u})$ son los valores esperados o medias aritméticas de las muestras disponibles y de la elevación en el área cubierta por el método de interpolación. Las diferentes variables de Kriging tienen como objetivo minimizar la varianza del error de estimación mediante la restricción

$$E\{Z^*(\mathbf{u}) - Z(\mathbf{u})\} = 0 \tag{8}$$

alcanzada a través de la asignación de pesos λ_α con base en la disimilaridad de los valores de las muestras $Z(\mathbf{u}_\alpha)$. La disimilaridad $\gamma(\mathbf{h})$ entre las muestras de elevación es establecida a través del semivariograma experimental definido como (Goovaerts, 1997:28):

$$\gamma(\mathbf{h}) = \frac{1}{2n(\mathbf{h})} \sum_{\alpha=1}^{n(\mathbf{h})} [z(\mathbf{u}_\alpha) - z(\mathbf{u}_\alpha + \mathbf{h})]^2 \tag{9}$$

donde $\mathbf{h}$ es la distancia entre las muestras $z(\mathbf{u}_\alpha)$ y $z(\mathbf{u}_\alpha + \mathbf{h})$ y n es el número de muestras separadas por la distancia $\mathbf{h}$.

Las variantes de Kriging empleadas en este estudio son Kriging Ordinario (OK, *Ordinary Kriging*) y Kriging con un modelo de tendencia (KT, *Kriging with a trend model*). Kriging Ordinario (OK) asume que la media $m(\mathbf{u})$ varía a través del área en estudio, pero es constante dentro de un área (o vecindad) local cubierta por el método de interpolación y, lo más importante, desconocida; por lo que es filtrada del método de interpolación estableciendo la suma total de los pesos de las muestras igual a 1 (*Ibid.*:133):

$$Z^*_{OK}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{OK}(\mathbf{u}) Z(\mathbf{u}_\alpha) + \left[1 - \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{OK}(\mathbf{u})\right] m(\mathbf{u})$$
$$= \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{OK}(\mathbf{u}) Z(\mathbf{u}_\alpha) \text{ con } \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{OK}(\mathbf{u}) = 1 \tag{10}$$

Kriging con un modelo de tendencia (KT) asume que $m(\mathbf{u})$ varía ligeramente en la vecindad local $W(\mathbf{u})$ del método de interpolación pero es desconocida. Dicha tendencia $m(\mathbf{u}')$ es modelada como una combinación linear de funciones $f_k(\mathbf{u})$ de las coordenadas geográficas K disponibles (Goovaerts, 1997: 140):

$$m(\mathbf{u}') = \sum_{k=0}^{K} a_k(\mathbf{u}') f_k(\mathbf{u}') \tag{11}$$
$$= \text{con } a_k(\mathbf{u}') \approx a_k \; \forall \; \mathbf{u}' \in W(\mathbf{u})$$

donde los coeficientes $a_k(\mathbf{u}')$ se consideran desconocidos y constantes para cada vecindad local. Dichos coeficientes son filtrados del algoritmo de interpolación estableciendo la suma total de los pesos igual a la tendencia $f_k(\mathbf{u}')$, (*Ibid.*)

$$Z^*_{KT}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{KT}(\mathbf{u}) Z(\mathbf{u}_\alpha) \tag{12}$$
$$\text{con } \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{KT} f_k(\mathbf{u}_\alpha) = f_k(\mathbf{u}) \quad k = 0, \ldots, K$$

CASO DE ESTUDIO

Área en estudio

La zona en estudio (Figura 1) se encuentra ubicada en el municipio de Yecapixtla, Morelos, México, y forma parte de un corredor para el cual se elaboró un estudio topográfico para la construcción de infraestructura por encargo de una empresa paraestatal mexicana; y se delimita por las coordenadas 18°50'50.36" N, 98°53'04.28" W; UTM WGS84 Zona 14 Norte: 512165 E, 2083938 N en la esquina superior izquierda, y 18°50'09.66" N, 98°52'21.60" W; UTM WGS84 Zona 14 Norte: 513415 E, 2082688 N en la esquina inferior derecha, formando un cuadrado de 1 250 x 1 250 m. El área cuenta con variada topografía (Figura 2) con un rango de elevaciones entre 1 492 y 1 586 m snmm. Es predominantemente plana al norte y montañosa al sur, y por lo tanto presenta las características apropiadas para evaluar el desempeño de diferentes métodos de interpolación en diferentes tipos de terreno.

Datos

Debido al estudio topográfico realizado previamente, se cuenta con una gran cantidad de datos para el área en estudio, incluyendo datos Lidar, fotografía aérea (Figura 2) y puntos de control recolectados

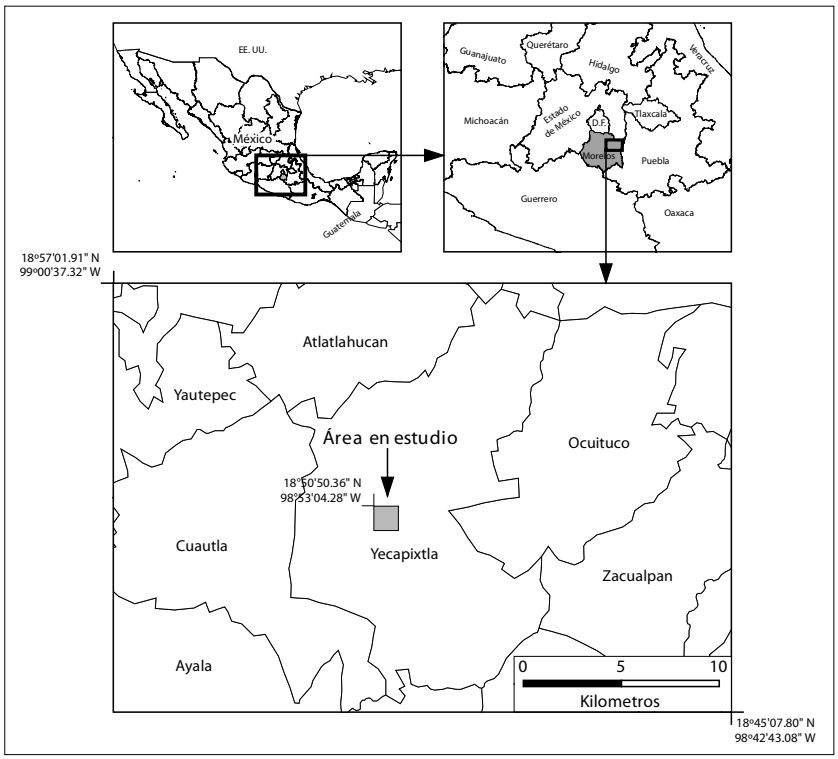

Figura 1. Localización del área en estudio.

con GPS cinemático de precisión. Para la realización de este estudio, se simuló la realización de un levantamiento topográfico para la generación de un MDE con tamaño de celda de 1 m (requerimiento de la paraestatal), para lo cual se solicitó a topógrafos con más de veinte años de experiencia en campo diseñaran dicho levantamiento. El diseño se basó en una combinación de muestras regulares y muestras representativas (Li *et al.*, 2005); e incluye una malla regular de puntos sobre el terreno cada 10 m y secciones cada 20 m en los rasgos hidrográficos (escurrimientos) más importantes y en la infraestructura carretera presente (Figura 2). En total, el levantamiento topográfico simulado se compone de 16 042 muestras de elevación tomadas de los datos Lidar disponibles.

Procesamiento

Las 16 042 muestras de elevación tomadas de los datos Lidar disponibles, fueron importadas en el paquete estadístico R (R Development Core Team, 2012) para ser procesadas con la librería gstat (Pebesma, 2004) y generar MDE con un tamaño de celda de 1 m usando los métodos de interpolación IDW (ecuación 1), OK (ecuación 10) y KT (ecuación 12). El MDE IDW fue generado usando 2 como exponente (ecuación 3) para ponderar la distancia de las doce muestras más cercanas al centro de cada celda de la cuadrícula y disminuir la presencia de patrones artificiales en el MDE. Los MDE OK y KT se generaron utilizando solo las 24 muestras más cercanas para permitir el cálculo variable de la media en la vecindad local de cada celda. El MDE correspondiente al método ANUDEM fue generado usando la versión del mismo implementada en ArcMap 10, sin forzar la definición de una red de drenaje, ya que dicha opción introdujo rasgos hidrográficos y topográficos artificiales debido a la alta densidad de datos (ESRI, 2011). Los cuatro MDE generados se muestran en la Figura 3

Una vez generados los MDE, la precisión de los mismos fue evaluada utilizando como referencia el MDE Lidar disponible para el área en estudio. El MDE Lidar tiene una precisión de 15 cm por lo que puede ser considerado como un MDE válido para evaluar la precisión de los MDE generados

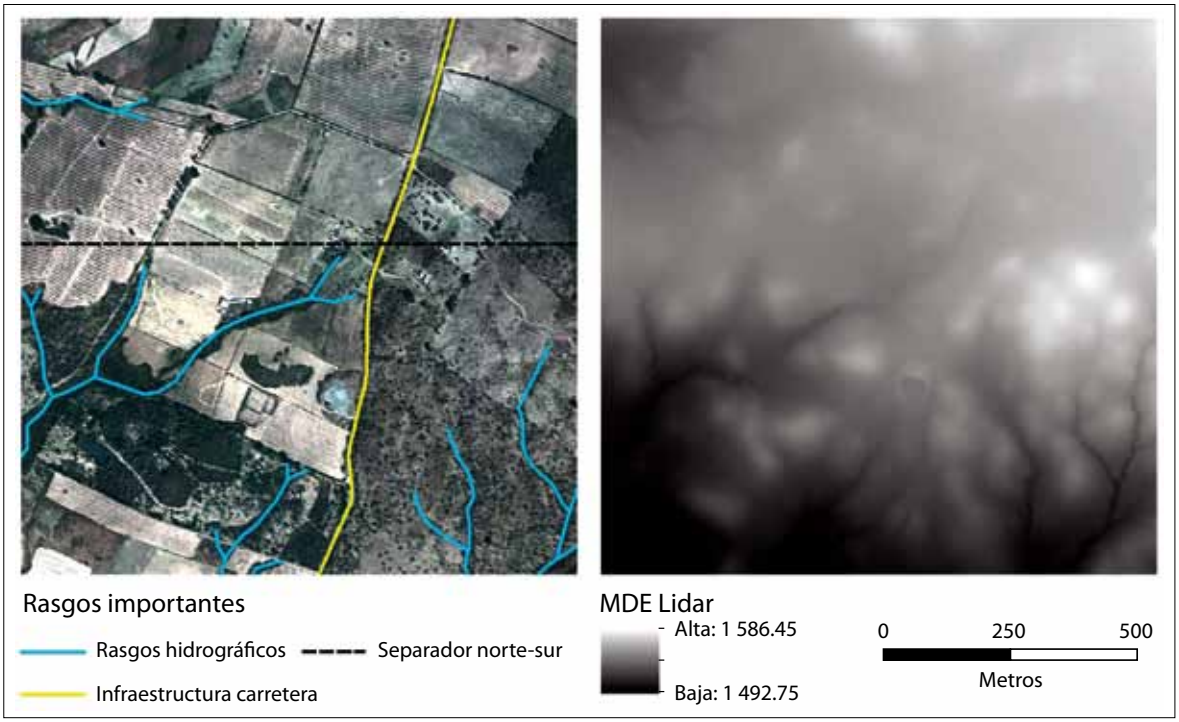

Figura 2. Fotografía aérea (izq.) y MDE Lidar (der.) de la zona en estudio. La fotografía aérea muestra los rasgos importantes para los cuales se definieron secciones cada 20 m y la separación entre zonas norte (plana) y sur (montañosa).

mediante interpolación (Fisher y Tate, 2006). La precisión o error de los MDE generados fue evaluada sustrayendo en cada celda la elevación reportada por el MDE Lidar de la elevación calculada utilizando interpolación. El resultado de esta operación fueron cuatro superficies de error (Figura 4) en las cuales valores positivos y negativos indican, respectivamente, sobreestimación y subestimación de la elevación por parte de los diferentes métodos de interpolación. Dichas superficies de error fueron utilizadas para generar estadísticas descriptivas de error, incluyendo el error medio cuadrático (EMC), de la zona en estudio completa (Tabla 1) y de las zonas norte (Tabla 2) y sur (Tabla 3). El EMC es la estadística comúnmente utilizada para reportar la precisión de MDE y definida como (*Ibid.*:470):

$$EMC = \sqrt{\frac{\sum (e_{DEM} - e_{REF})^2}{n}} \qquad (13)$$

donde e_{DEM} es la elevación de cada celda del MDE, e_{REF} es la elevación del MDE utilizado como referencia para evaluar la precisión y n el número de celdas evaluadas.

Resultados

Los diferentes métodos de interpolación determinísticos (IDW y ANUDEM) y geoestadísticos (OK y KT) usados en este estudio generaron MDE que gráficamente representan la mayoría de las características topográficas presentes en el área en estudio. No obstante, los MDE IDW (Figura 3a) y ANUDEM (Figura 3b) muestran patrones artificiales, que pueden ser clasificados como errores sistemáticos, en las celdas cercanas a las muestras de elevación. Dichos patrones se pueden apreciar fácilmente en la zona sur de las superficies de error correspondientes a los métodos de interpolación determinísticos (Figura 4a y b). En contraste, los MDE OK (Figura 3c) y KT (Figura 3d) no presentan patrones artificiales. Sin embargo, el MDE OK representa los rasgos topográficos del área de manera demasiado suave, lo cual también puede ser considerado como error sistemático; e inclusive

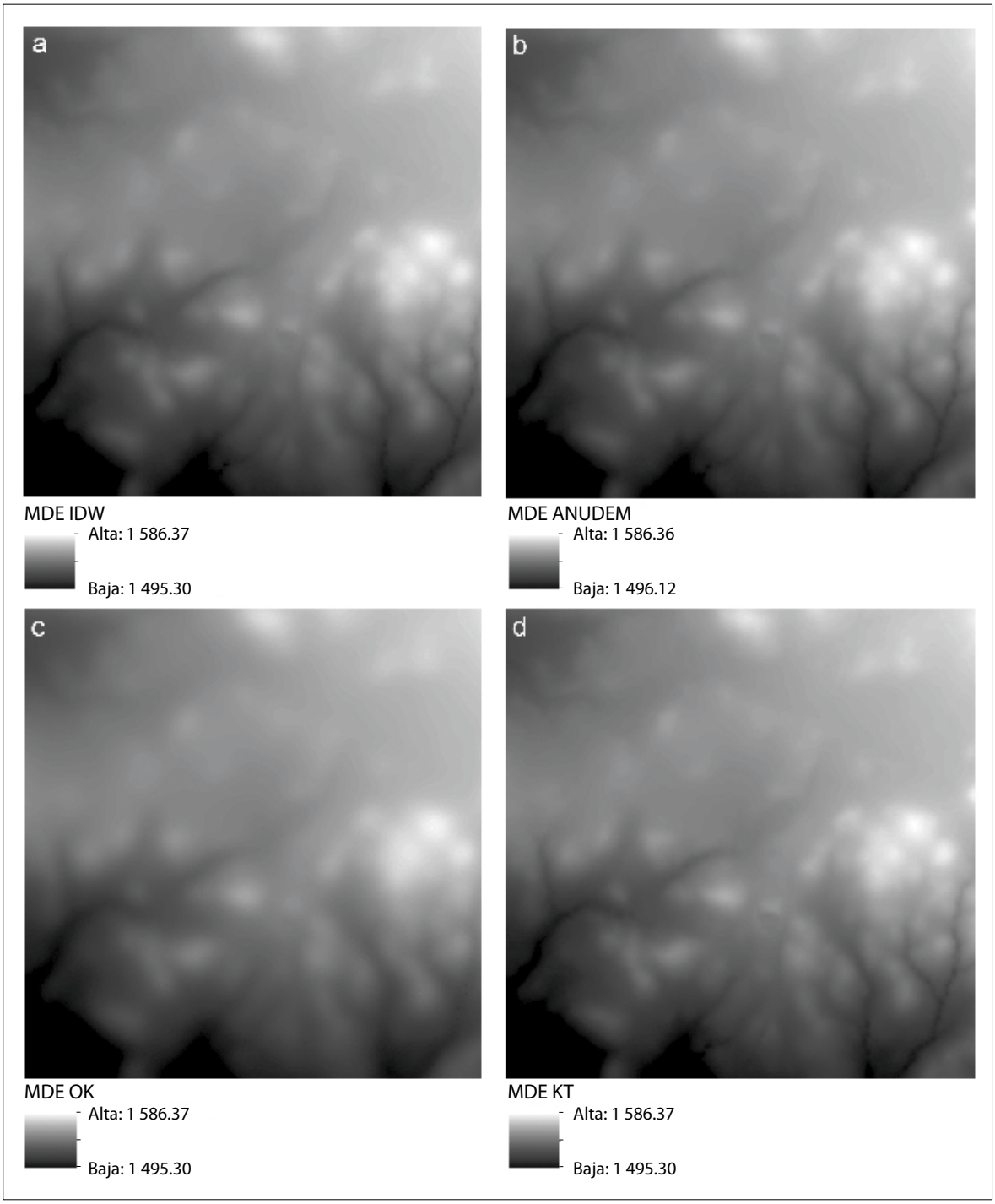

Figura 3. MDE generados mediante métodos de interpolación. a. MDE IDW. b. MDE ANUDEM. c. MDE OK. d. MDE KT.

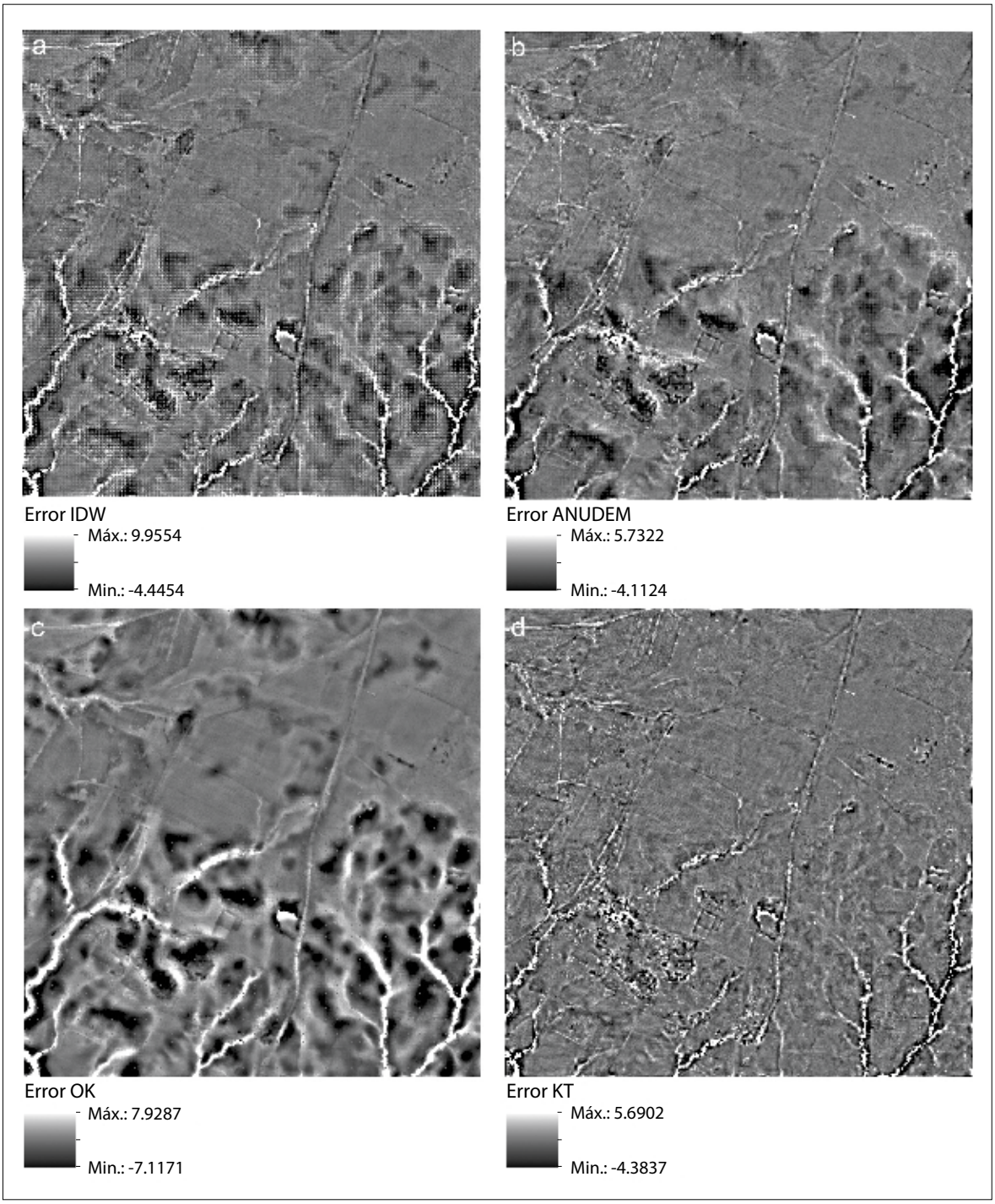

Figura 4. Superficies de error de los MDE mostrados en la Figura 3. a. Error MDE IDW. b. Error MDE ANUDEM. c. Error MDE OK. d. Error MDE KT.

Tabla 1. Estadísticas descriptivas del error en los MDE generados mediante interpolación

MDE	Estadísticas de error (m)								
	EMC	Media	Desv. Est.	Min	P 0.025	P 0.250	P 0.750	P 0.975	Max
IDW	0.2565	-0.0150	0.2560	-4.4454	-0.5292	-0.0983	0.0789	0.4073	5.9554
ANUDEM	0.2377	-0.0443	0.2335	-4.1124	-0.5090	-0.1099	0.0353	0.3134	5.7322
OK	0.6145	-0.0189	0.6142	-7.1171	-1.3877	-0.2201	0.2114	1.1160	7.9287
KT	0.1804	-0.0026	0.1805	-4.3837	-0.2983	-0.0566	0.0507	0.2797	5.6902

Tabla 2. Estadísticas descriptivas del error en la zona norte (plana) de los MDE generados

MDE	Estadísticas de error (m)								
	EMC	Media	Desv. Est.	Min	P 0.025	P 0.250	P 0.750	P 0.975	Max
IDW	0.1291	-0.0037	0.1291	-3.2803	-0.2859	-0.0542	0.0574	0.2285	1.3016
ANUDEM	0.1078	-0.0168	0.1065	-2.4119	-0.2411	-0.0621	0.0326	0.1855	1.2135
OK	0.2752	0.0072	0.2751	-4.7847	-0.6849	-0.0820	0.1366	0.4930	1.7748
KT	0.0904	0.0004	0.0904	-1.6843	-0.1843	-0.0417	0.0428	0.1843	1.1865

Tabla 3. Estadísticas descriptivas del error en la zona sur (montañosa) de los MDE generados

MDE	Estadísticas de error (m)								
	EMC	Media	Desv. Est.	Min	P 0.025	P 0.250	P 0.750	P 0.975	Max
IDW	0.3173	-0.0230	0.3165	-4.4454	-0.6368	-0.1497	0.1057	0.5306	5.9554
ANUDEM	0.2973	-0.0639	0.2904	-4.1124	-0.6394	-0.1599	0.0386	0.4220	5.7322
OK	0.7697	-0.0375	0.7687	-7.1171	-1.6245	-0.4025	0.3151	1.5537	7.9287
KT	0.2235	-0.0048	0.2234	-4.3837	-0.3823	-0.0708	0.0583	0.3689	5.6902

elimina algunos rasgos secundarios del área, dando como resultado la presencia de errores importantes aun en la parte plana del área en estudio (Figura 4c). El MDE KT (Figura 3d) no presenta patrones visibles como se puede confirmar en la superficie de error correspondiente (Figura 4d), donde se muestra que los errores del MDE KT se distribuyen de forma uniforme en el área de estudio.

La evaluación estadística del error en los MDE generados (Tabla 1) confirma la evaluación visual de los MDE. Todos los MDE generados representan satisfactoriamente la superficie terrestre y, en consecuencia, tienen un EMC bajo, con medias de error cercanas a 0 y, por lo tanto, desviaciones estándar muy cercanas al EMC (Li *et al.*, 2005). La presencia de errores sistemáticos en los MDE IDW, ANUDEM y OK no es evidente en las estadísticas presentadas en la Tabla 1, por lo que es necesario recurrir a las estadísticas de error generadas para los diferentes tipos de terreno presentes en el área en estudio. En la zona norte del área, donde el terreno es predominantemente plano, las estadísticas de error (Tabla 2), de igual forma, indican que los cuatro métodos de interpolación son apropiados para generar MDE. En esta zona, la diferencia entre la magnitud de los errores mínimos y máximos indica la existencia de errores sistemáticos (subestimación) en los MDE IDW, ANUDEM y OK. En la zona montañosa (sur) del área en estudio, la presencia de errores sistemáticos es más evidente con errores medios más alejados del 0 (Tabla 3), con excepción de la media de error en el MDE KT.

En términos de precisión absoluta, las estadísticas de error indican que KT es el método de interpolación más preciso de los cuatro métodos evaluados en este estudio, tanto en zonas con topografía variada (Tabla 1) como en zonas predominantemente planas (Tabla 2) y montañosas (Tabla 3). En el área en estudio completa el EMC de KT es cercano al de IDW y ANUDEM, no obstante, la distribución de los errores en el MDE KT (con un 75% de los errores entre -5.66 y 5.07 cm y un 95% de los mismos entre -29.83 y 27.97 cm) permite establecer que KT es más confiable que el resto de los métodos de interpolación, en los cuales la distribución de errores no es uniforme (Tabla 1). En la misma zona, OK generó el MDE con el EMC más alto debido principalmente al efecto suavizante del método. En la zona norte del área en estudio la precisión del MDE KT es similar a la de los MDE IDW y ANUDEM, no así a la del MDE OK debido al efecto mencionado anteriormente. En la zona sur las estadísticas de error (Tabla 3) muestran de igual forma que el MDE KT tiene una mejor precisión que la del resto de los MDE; con un EMC más bajo y una distribución uniforme de los errores, en contraste con el resto de los MDE donde la influencia de errores sistemáticos es evidente.

CONCLUSIONES

Los MDE contienen errores de forma inherente debido a diferentes factores. En este trabajo se evaluó y comparó la precisión de MDE generados utilizando métodos de interpolación determinísticos (IDW y ANUDEM) y probabilísticos (OK y KT), ya que existe cierta discrepancia entre la comunidad científica acerca de la idoneidad de diferentes métodos de interpolación para la generación de MDE. Los resultados del caso de estudio aquí presentado sugieren que KT genera MDE más precisos que IDW, ANUDEM y OK. No obstante, no es posible establecer qué métodos probabilísticos de interpolación generan MDE más precisos que métodos determinísticos, ya que el MDE con el EMC más alto fue generando usando OK (uno de los métodos probabilísticos).

Asimismo, aun cuando KT generó el MDE más representativo gráficamente (sin errores sistemáticos) y con la mejor precisión estadística en los diferentes tipos de terreno evaluados y cuando estos resultados concuerdan con aquellos de Lloyd y Atkinson (2002; 2006), Erdogan (2009) y Guo *et al.* (2010), es necesaria la elaboración de otros estudios que confirmen la superioridad de KT sobre otros métodos de interpolación. Lo anterior debido a que los resultados del caso de estudio, al igual que los de los otros estudios citados anteriormente, discrepan con los resultados de Bishop y McBratney (2002), Su y Bork (2006) y Bater y Coops (2009).

Es importante observar que tanto el caso de estudio aquí presentado, como aquéllos con los que se contrastan los resultados del mismo, fueron elaborados utilizando una gran densidad de muestras de elevación para la generación de MDE de alta resolución con tamaños de celda menores a 5 m. Por lo tanto, para la generalización de los resultados aquí presentados se requiere de la elaboración de otros estudios que consideren combinaciones de bajas densidades de datos y bajas resoluciones.

Finalmente, es también importante mencionar que aun cuando el MDE KT tiene un EMC muy cercano al del MDE Lidar, 18 y 15 cm, respectivamente, el MDE Lidar representa una mayor cantidad de características sobre el terreno en comparación con los MDE generados mediante interpolación. Esto se puede observar gráficamente en las superficies de error (Figura 4) descritas anteriormente, donde se pueden observar en las cuatro superficies de error los rasgos no representadas por los MDE generados.

REFERENCIAS

Anderson, E. S., J. A. Thompson, D. A. Crouse and R. E. Austin (2006), "Horizontal resolution and data density effects on remotely sensed LIDAR-based DEM", *Geoderma*, no. 132, pp. 406-415.

Anderson, J. M. and E. M. Mikhail (1998), *Surveying: theory and practice*, 7th ed., McGraw-Hill, Nueva York.

Bater, C. W. and N. C. Coops (2009), "Evaluating error associated with Lidar-derived DEM interpolation", *Computers & Geosciences*, no. 35, pp. 289-300.

Bishop, T. F. A. and A. B. McBratney (2002), "Creating field extent digital elevation models for precision agriculture", *Precision Agriculture*, no. 3, pp. 37-46.

Bosque Sendra, J. (2000), *Sistemas de Información Geográfica*, Ediciones Rialp, Madrid.

Burrough, P. A. and R. A. McDonnell (1998), *Principles of Geographical Information Systems*, Oxford University Press, London.

Erdogan, S. (2009), "A comparison of interpolation methods for producing digital elevation models at the field scale", *Earth Surface Processes and Landforms*, no. 34, pp. 366-376.

ESRI (2011), "Cómo funciona de topo a raster", ARCGIS Resource Center, Environmental Systems Research Institute [http://help.arcgis.com/es/arcgisdesktop/10.0/help/index.html#//009z0000007m000000: 12 de junio de 2012].

Fisher, P. F. and N. J. Tate (2006), "Causes and consequences of error in digital elevation models", *Progress in Physical Geography*, no. 30, pp. 467-489.

Goovaerts, P. (1997), *Geostatistics for natural resources evaluation*, Oxford University Press, New York.

Guo, Q., W. Li, H. Yu and O. Álvarez (2010), "Effects of topographic variability and Lidar sampling density on several DEM interpolation methods", *Photogrammetric Engineering & Remote Sensing*, no. 76, pp. 701-712.

Hengl, T. (2006), "Finding the right pixel size", *Computers & Geosciences*, no. 32, pp. 1283-1298.

Hutchinson, M. F. (1988), "Calculation of hydrologically sound digital elevation models", in *Proceedings of the Third International Symposium on Spatial Data Handling*, 17-19 August, Sydney, Australia, International Geographical Union, Ohio, pp. 117-133.

Hutchinson, M. F. (1996), "A locally adaptive approach to the interpolation of digital elevation models", in *Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modeling*, National Center for Geographic Information and Analysis, 21-25 January, Nuevo México.

Hutchinson, M. F. (2008), "Adding the Z dimension", in Wilson, J. P. y A. S. Fotheringham (coords.), *The Handbook of Geographic Information Science*, Blackwell, Oxford, pp. 144-168.

INEGI (2002), *Tratamiento de errores en levantamientos topográficos*, Instituto Nacional de Estadística, Geografía e Informática, Aguascalientes, México.

Kumler, M. P. (1994), "An intensive comparison of Triangulated Irregular Networks (TINs) and Digital Elevation Models (DEMs)", *Cartographica*, no. 31, pp. 1-99.

Li, Z., Q. Zhu and C. Gold (2005), *Digital terrain modeling. Principles and methodology*, CRC Press, Florida.

Lloyd, C. D. and P. M. Atkinson (2002), "Deriving DSMs from LIDAR data with Kriging", *International Journal of Remote Sensing*, no. 23, pp. 2519-2524.

Lloyd, C. D. and P. M. Atkinson (2006), "Deriving ground surface digital elevation models from LIDAR data with geostatistics", *International Journal of Geographical Information Science*, no. 20, pp. 535-563.

Longley, P. A., M. F. Goodchild, D. J. Maguire and D.W. Rhind (2005), *Geographical information systems and science*. Wiley, Chichester, London.

López, C. (2000), "Improving the elevation accuracy of digital elevation models: a comparison of some error detection procedures", *Transactions in GIS*, no. 4, pp. 43-64.

Maune, D. F., S. M. Kopp, C. A. Crawford and C. E. Zervas (2001), "Introduction", in Maune, D. F. (coord.), *Digital elevation model technologies and applications: the DEM users manual,* The American Society for Photogrammetry and Remote Sensing, Maryland, pp. 1-34.

Miller, C. L. and R. A. Laflamme (1958), "The digital terrain model - theory and application", *Photogrammetric Engineering*, no. 24, pp. 433-442.

Morillo, J., J. Pozo, F. Pérez, M. C. Rodríguez, F. Gordillo y J. Rebollo (2002), "Análisis de calidad de un modelo digital de elevaciones generado con distintas técnicas de interpolación", en *XIV Congreso Internacional de Ingeniería Gráfica*, 5-7 de junio, Santander, España [http://departamentos.unican.es/digteg/ingegraf/cd/ponencias/11.pdf: 12 de junio de 2012].

O'Sullivan, D. and D. J. Unwin (2003), *Geographic Information Analysis*, Wiley, Nueva Jersey.

Pebesma, E. J. (2004), "Multivariable geostatistics in S: the gstat package", *Computers & Geosciences*, no. 30, pp. 683-691.

Pérez Vega, A. y J.-F. Mas (2009), "Evaluación de los errores de modelos digitales de elevación obtenidos por cuatro métodos de interpolación", *Investigaciones Geográficas, Boletín*, núm. 69, Instituto de Geografía, UNAM, México, pp. 53-67.

R Development Core Team (2012) "R: a language and environment for statistical computing", *R Foundation for Statistical Computing*, Vienna, Austria [http://www.R-project.org/].

Reuter, H. I., A. Nelson and A. Jarvis (2007), "An evaluation of void-filling interpolation methods for SRTM data", *International Journal of Geographical Information Science*, no. 21, pp. 983-1008.

Smith, S. L., D. A. Holland and P. A. Longley (2005), "Quantifying interpolation errors in urban airborne laser scanning models", *Geographical Analysis*, no. 37, pp. 200-224.

Su, J. and E. Bork (2006), "Influence of vegetation, slope and Lidar sampling angle on DEM accuracy", *Photogrammetric Engineering and Remote Sensing*, no. 72, pp. 1265-1274.

USGS (1997), *Standards for digital elevation models*, United States Geological Survey, Department of the Interior, Washington D.C.

Wise, S. (2000), "Assessing the quality for hydrological applications of digital elevation models derived from contours", *Hydrological Processes*, no. 14, pp. 1909-1929.

Wise, S. (2011), "Cross-validation as a means of investigating DEM interpolation error", *Computers & Geosciences*, no. 37, pp. 978-991.